Manuscript prepared for J. Name
with version 2.3 of the LaTeX class copernicus.cls.
Date: 17 November 2012

Supplementary material to: Top-down estimation of carbon monoxide emissions from the Mexico Megacity based on FTIR measurements from ground and space

W. Stremme[1], M. Grutter[1], C. Rivera[1], A. Bezanilla[1], A.R. Garcia[1], I. Ortega[1], M. George[2], C. Clerbaux[2], P. Coheur[2], D. Hurtmans[2], J.W. Hannigan[3], and M.T. Coffey[3]

[1]Centro de Ciencias de la Atmósfera, Universidad Nacional Autónoma de México, México City
[2]UPMC Univ. Paris 06; Univ. Versailles St-Quentin; CNRS/INSU, LATMOS-IPSL, Paris, France
[3]Atmospheric Chemistry Division, NCAR, Boulder CO, United States

Introduction

The ground and space based measurements and the technique used to determine an annual top down emission estimation of Carbon Monoxide (CO) on a local scale are described in the main article "Top-down estimation of carbon monoxide emissions from the Mexico Megacity based on FTIR measurements from ground and space". The method, using columnar measurements is in principal straight forward, however various technical details of the implementation and the parameters required and their quantitative impact on the result are evaluated in this supplement.

Two novel techniques are developed for the purpose of top down emission estimation: i) The reconstruction of the annual mean CO column distribution of the strongly inhomogeneous Mexico City Metropolitan Area (MCMA) from the satellite based measurements and ii) the technique to compare and combine measurements from space and ground above the inhomogeneous area.

While the main technical procedure of the reconstruction of the annual distribution is described in the main article, Section 1 of the supplement addresses the "averaging kernel" (AK) as a diagnostic tool and also describes an extended error analysis for the reconstructed column density distribution. The AK is developed in analogy to Optimal Estimation Theory and gives the sensitivity of the resultant state vector to the real state where bottom up and modelled estimations can aid in defining the emissions. Confidence in the AK can then be used too iteratively tune the retrieval. The use of OE theory and the AK of the 2D- field of column densities along with the definition of spatial resolution of the annual mean column density derived from long term measurements is new and

therefore described in this supplementary part.

Section 2 of the supplement addresses in more detail the inter-comparison of the measurements. Satellite validation of measurements over a footprint with inhomogeneous column density is a precondition for combining them in the top down emission estimation, an actual open topic itself, and it is therefore addressed in more detail in this section. Correlative long term wind measurements are used to construct an operator which accounts for local transport and comparison when the measurements are not strongly coincident. A comparison with a conventional validation study with this alternative method is presented and its technical procedure is described in detail.

1 IASI CO distribution: diagnostics and errors

Error estimation and characterisation of a retrieval which uses *apriori* information and constraints has been for more than 20 years integral in remote sensing techniques. The importance of the diagnostics using averaging kernels increases a) with measurements of higher vertical (spatial) resolution are presented, for example profiles instead of total columns and b) when a quantitative result is presented instead a relative behaviour. The analogy of the AK as the most important diagnostic tool for a profile retrieval to a reconstruction of horizontal distribution of annual mean column densities is straight forward. The utility of this diagnostic tool for the application here is a) illustrated in the calculated spatial resolution Sect. 1.1 and is therefore interesting for the reconstruction of urban climatologies based on remote sensing data. And b) provides feedback in order to adjust the constraint as shown in Sect. 1.2 . The total content averaging kernel for the MCMA (Sect. 1.2) illustrates the ability to calculate the annual total content of CO in the Mexico City mixing layer from the reconstructed distribution. The total content is here defined as the integral of the partial mixing layer column of CO over the MCMA.

1.1 Spatial resolution of IASI CO

Information contained in an atmospheric gas retrieval using IR-spectroscopy and Optimal Estimation inversion theory is commonly quantified with the degrees of freedom of signal (DOFS) which gives the number of independent pieces of information in the derived state vector. In the retrieval of the spatial distribution of CO the DOFS will give a measure of the spatial resolution of the average column densities of CO. This resolution is in part determined by the strength of the a priori constraint. In profile retrievals the DOFS is calculated as the trace of the profile averaging kernels and the same can be done for the reconstruction of spatial distribution. The averaging kernel $\mathbf{A}$ for CO distribution is calculated according to the formulas of Rodgers (2000), while measurement-noise matrix $\mathbf{S}_e$ is chosen to be diagonal and uniform, so that the averaging kernel for CO distribution $\mathbf{A}$ only depends on $\mathbf{K}$ and $\mathbf{R}$, where $\mathbf{R}$ is the regularisation matrix which represents the constraint of

the *apriori*.

$$\mathbf{A} = (\mathbf{K^T K} + \mathbf{R})^{-1} \mathbf{K^T K} \tag{1}$$

The forward model is linear, thus it is given by its Jacobian $\mathbf{K}$, and describes the distribution with two variables (background and mixing layer) in each grid-point, however the spatial variability of the background layer is low. Therefore the background is constrained by a rather strong smoothness constraint, does not play an important role for resolution and is ignored in the following description and only the block of the averaging kernel related with the mixing layer column of CO is considered. Even if we consider a 2-D distribution we treat it as a 1-D vector and in the following equation (main article equation 6)

$$\boldsymbol{TC}^{day}_{IASI} = \sum_{i}^{N_{footprint}} (\frac{AB}{N_{footprint}} \cdot \hat{\lambda}^i_1 + \frac{AM}{N_{footprint}} \hat{\lambda}^i_2) + \boldsymbol{\epsilon} \tag{2}$$

the CO mixing layer column amount on the position $i(l, m)$ of the gridded area (l,m are the corresponding coordinates of the grid) is described by λ^i_2.

For direct comparison the ground based measurement site should be in the field of view of the IASI instrument where this field of view determines the horizontal resolution, however, in the calculation of an average horizontal distribution, the field of view size is a parameter of the forward model. The total horizontal information content is given by trace(**A**). The information per grid cell is given by each diagonal element such that $\mathbf{A}_{k,k}/\Delta S$ represents the information density in DOFS per unit area, where ΔS is the area corresponding to each grid point, such that the integration over the whole area S results in the DOFS=Trace(**A**) or DOFS = $\int \frac{A_{k,k}}{\Delta S} dS$. The horizontal resolution of the average CO column density of IASI in the grid point k(l,m) can be described by the square root of the inverse of the information density ($\sqrt{\Delta S/\mathbf{A}}_{k,k}$) and has the unit "km" and should be largely independent of the grid. The information density is shown in the Figure 1 for the total column of CO which is the sum of both mixing layer and background CO partial columns. By tuning increasing the constraint for the background CO, the total information density is increasingly provided by the mixing layer CO partial column. The use of the diagonal of the Averaging Kernel and the definition of a resolution is motivated through an analogy to profile retrieval for limb sounding (e.g. Funke et al., 2009).

In the center of the Mexico City Basin the information density of IASI CO is about 0.02 DOFS per km^2 which reflects an horizontal resolution (as it is defined here) of around 7 km ($\approx 1/\sqrt{0.02\ \mathrm{DOFS/km^2}}$). This is below the NADIR footprint of IASI of 12 km.

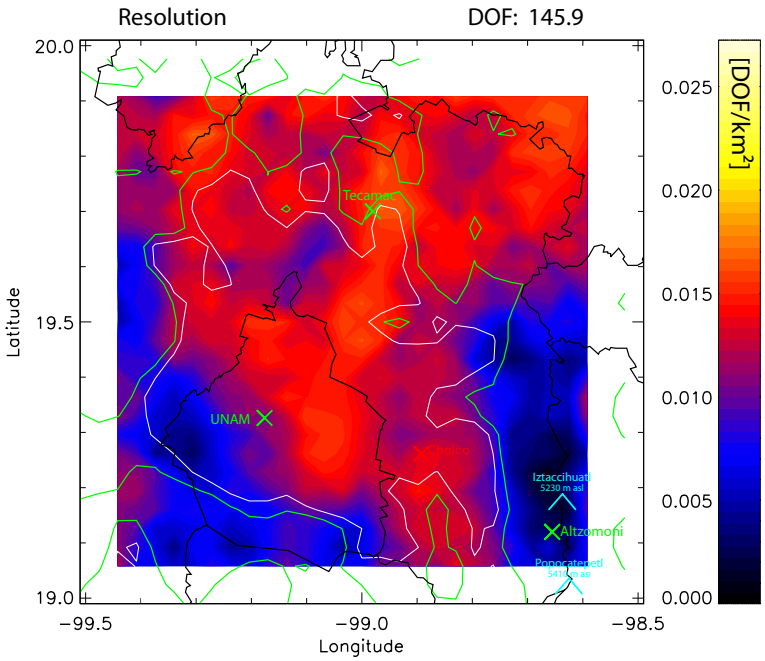

Fig. 1. Average Kernel Diagonal. The resolution of the retrieved horizontal gas distribution can be estimated by the diagonal of an Averaging Kernel matrix, which is calculated analogously to vertical gas distribution retrievals. To get a independent quantity (DOFS=1) near CCA-UNAM, the value , 0.0156 DOFS/km^2 has to be multiplied by an area of 8km $\times$ 8km and defines somehow the horizontal resolution. There occurs a retrieval-artifact near the edges, as the information of measurements is just distributed to few grid-points, the information density is higher (up to a factor of 2). The area is mostly outside of the metropolitan area and does not affect the results and therefore it is not ploted.

1.2 Total content Averaging Kernel for IASI CO

The strength of the horizontal background and mixing layer constraints are tuned so that the background CO partial column is nearly homogeneous and the resulting distribution of the mixing layer CO column matches roughly with the topography (mountains) and population distributions (suburban municipality) of Mexico City and the resulting MCMA is still a connected area. It is possible to obtain similar formed distributions and resolutions using other constraints (diagonal or L_1-Thikonov) and the type of constraint may be optimized for different purposes (Steck, 2002). The calculated CO emission of the Metropolitan area depends on the kind and strength of the constraint and it has to be checked whether the total content of CO in the MCMA is correctly quantified. Which anomalies are underestimated, overestimated or correctly determined by the reconstruction can be determined by the total content MCMA Averaging Kernel and is the topic of this section.

Following the analogy to profile retrievals and the total column operator of Rodgers and Connor (2003) we define the total content MCMA operator $\boldsymbol{g}^t_{MAMC}$. Iteratively the constraint has been optimized until the total content averaging kernel $A_{MCMA} = \boldsymbol{g}^t_{MCMA}\mathbf{A}$ (Fig. 2) has a nearly ideal form:

$$A^{ideal}_{MCMA} = \left\{ \begin{array}{l} 1 \ \text{inside MCMA} \\ 0 \ \text{outside MCMA} \end{array} \right\}$$

In general the retrieval product together with an averaging kernel matrix provide the data for further geophysical interpretation or comparison.

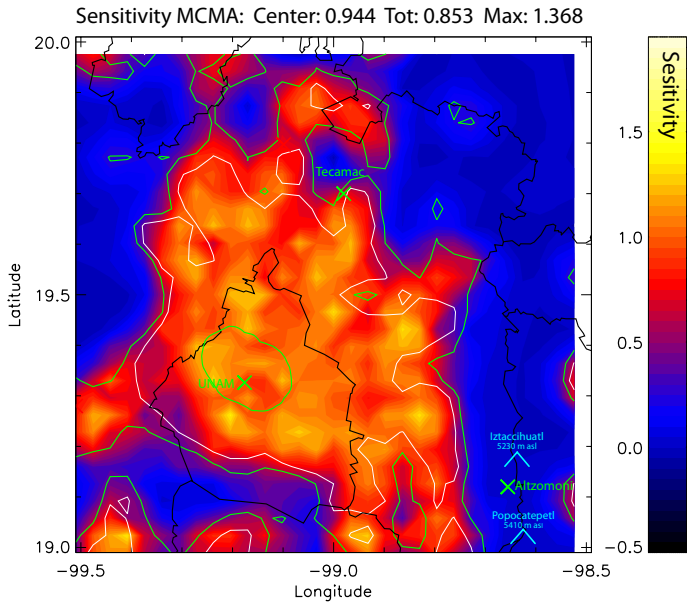

Fig. 2. Sensitivity of the total content of CO in the Mexico City Metropolitan Area. Diagnostic tool $A_{MCMA} = \boldsymbol{g}^t_{MCMA}\mathbf{A}$: Mexico City Metropolitan Area - total content - averaging kernel is calculated analogously to the total or a partial column averaging kernel. It shows the sensitivity to measure CO molecules in the MCMA. The area MCMA is shown by the outer green contour line. The white contour line includes an area which has more than 1.65E18 molec/cm^2. The area between the white and the outer green contour lines might indicate where the area change from urbanized to rural. The inner circle around UNAM (green cross) represents the area in which the CO measured in the vertical column between 11:15 and 13:15 at UNAM was with a 85% probability located two hours before at MetOp-A overpass time.

1.3 Random error in the average IASI-CO distribution

Analogous to the error estimation for atmospheric profile retrievals by Rodgers (1990), the random error or measurement error for the average CO distribution is described as a covariance Matrix

$$\mathbf{S}_m = \cdot\mathbf{D}\cdot S_y \cdot \mathbf{D} \tag{3}$$

where the square root of its diagonal is a vector which describes the error in each grid point and err_{MCMA}, the error of the number of molecules in the MCMA, is described by the following expression.

$$err_{MCMA} = \sqrt{\boldsymbol{g}^t \cdot \mathbf{D} \cdot \mathbf{S}_y \cdot \mathbf{D} \cdot \boldsymbol{g}} \tag{4}$$

The covariance matrix of the measurement noise S_y has to be estimated.

However there is only one residual $\boldsymbol{dy}$ available which could be used to estimate $\mathbf{S}_y$.

$$\boldsymbol{dy} = (y - Kx) \tag{5}$$

The residual contains apart from random noise also a time dependent structure resulting from the seasonal cycle of CO, which is not represented in the forward model of this study and might be included in the measurement error using a generalized $\mathbf{S}_y$ in eq. 3 (von Clarmann et al., 2001). The random noise includes also day to day variation of the columnar CO which originates from variation in traffic or meteorological conditions or other events such as regional fires. As both y and x (in Eq. 5) are describing columnar CO, we can calculate an uncertainty from the residual in the y-space (CO column) and then map the error pattern to x-space (on the grid) with Equation 3 and using the **Gain**-matrix ($\mathbf{D}$).

$$\mathbf{D} = (K^T K + R)^{-1} K^T \tag{6}$$

Unfortunately the residual is on the one hand the only information source for the noise pattern and on the other a result of the retrieval and is therefore not an independent noise pattern, even it should reflect the systematic and random patterns which are expected.

The in time mirrored residual $\boldsymbol{dy}_{mirrored}$ is used as a typical residual to calculate $\mathbf{S}_y = \boldsymbol{dy}^t_{mirrored}\boldsymbol{dy}_{mirrored}$. It is used as input for the gain matrix and contains all statistical properties of the noise, random noise and systematic features, but it is not a direct product of the retrieval. The matrix $\mathbf{S}_y$ is calculated from a single vector and not from an ensemble, therefore Equation 4 could be simplified:

$$\boldsymbol{\epsilon}_{mirror} = \mathbf{D} \cdot \boldsymbol{dy}_{mirrored} \tag{7}$$

$$error_{IASI} = RMS(\boldsymbol{\epsilon}_{mirror}) \tag{8}$$

$$error_{MCMA} = \boldsymbol{g}^t \cdot \boldsymbol{\epsilon}_{mirror} \tag{9}$$

The random errors in each grid point are estimated to be around 2% of the mean CO column and the error in the total content of CO, the mixing layer partial CO columns integrated over the MCMA, results in a relative error of 8%.

As the error is calculated from a single residual and the resulting error is an estimation for a typical error, though it is not the quantified statistical error and therefore an alternative path for error estimation was used in the main article. The difference of two retrievals with first a slightly too strong and second a slightly too low constraint contains both a contribution of the smoothing error and of the measurement error.

2 Comparison and validation details

In this section the columns retrieved at the UNAM station from solar absorption FTIR measurements are compared with the IASI satellite product for CO. First, only mean values inside and outside the MCMA are compared from the entire data set. Then, individual measurements are compared with certain coincidence criteria (shown in the left columns in Tabs. 2 and 3) and the results are presented in right columns in the same tables below. As the comparison shows that a sufficient strict coincident criteria leads to very few coincident measurement events thus the result does not allow for conclusions about the bias, another approach is used in the main article and the details how the correlative wind measurements in the Mexican City Basin are used for this purpose is given in Sect. 2.3.

2.1 Comparison of mean values

For a comparison of the average background values, outside the MCMA like at the Tecámac site, no significant diurnal cycle is expected, so that the mean value of the IASI distribution of 1.65×10^{18} molec/cm^2 might be compared with 1.72×10^{18} molec/cm^2 the mean value of the measurements taken during March 2006 as part of the MILAGRO field campaign. The measurements agree quite well considering the Tecámac measurements were not restricted to the overflight times and the campaign of MILAGRO was in spring, the season for which a slightly higher column of CO is expected. The difference of 0.07×10^{18} molec/cm^2 is smaller than the standard deviation of the measurements performed within that one month 0.13×10^{18} molec/cm^2, and the expected amplitude of the seasonal cycle ($> 0.1 \times 10^{18}$ molec./cm^2) and therefore no significant bias is detected in the range of its uncertainty.

For a comparison of the measurements done at the UNAM within the MCMA, the diurnal cycle has to be taken into account. The average of all columns measured at around 10:19 LT on days with low ventilation is $(2.59 \pm 0.13) \times 10^{18}$ molec/cm^2, are significantly higher than the reconstructed IASI measurements value of 2.13×10^{18} molec/cm^2 at the location of the ground measurements. This difference of 20% (0.46×10^{18} molec/cm^2) is significant and might originate for different reasons. In

Table 1. Results for averaged measurements of the three instruments.

	CCA-UNAM	IASI		Tecámac
	ground based	space based		ground based
raw data	CO - total columns in [molec./cm^2]			
	selected	day	night	all
Average	3.06e+18	1.47e18	1.35e18	1.72e+18
Standard deviation	0.76e+18	0.34e+18	0.22e+18	0.22e+18
AK corrected [a]	-	1.55e+18	1.49e+18	-

a) Retrieved columns have been corrected according to a MLH of 1.5 km (see text).

the morning overflight time of IASI the sun zenith angle is quite high and the CO distribution is not homogeneous. A comparison has to be done with care as there is typically more contamination to the east of UNAM. Therefore, the value should be compared with a slightly displaced value. In Figure 6 (main article) the annual mean of the displaced location is shown by the smaller cross labelled as "10:19" and has the value of 2.26×10^{18} molec./cm^2. The displacement is calculated assuming a mixing (or residual) layer height of 2.5 km. So that the slanted sun light might have crossed a layer with an elevated amount of CO at that time.

Closer to mid-day, the CO in the well mixed residual layer has been disconnected from the surface CO emission and disappears slowly, while the column in the new mixing layer spans a very small corresponding horizontal area which is sufficiently mixed and therefore measurements represent a very local airmass. It will be shown in the next section that the emission varies strongly depending on the footprint size assumption. The IASI-distribution, however, automatically averages the area (horizontal resolution of around 8 km around UNAM, Fig.1), so that the local effect of emissions from sources outside the campus are reduced for the mean value of the IASI data but not for the UNAM column measurements.

2.2 Direct Comparison and Validation of coincident measurements

A direct comparison between the ground-based and satellite measurements is performed for different coincidence criteria. Table 2 shows the comparisons using only the dataset of weekdays with low ventilation, which are used for the top-down emission estimation. The columns at UNAM are calculated from the average of measurements taken within one hour around the IASI overpass time on a specific day. The corresponding IASI data falling within a circle of a given radius (first column in Table 2) around UNAM campus is taken. If more than one IASI measurement is available then the average is taken. Due to the different instrumental sensitivities, their averaging kernels are used. For the comparison a "true profile shape" is estimated for each ground-based measurement. This profile uses a constant volume mixing ratio (VMR) in the bottom layer up to the altitude if the mixing layer

height (MLH). The MLH is a reconstruction from the column measurement and the surface concentration of CO as described by Stremme et al. (2009). Due to the optimization of the CO column retrieval for mixing layer pollution, the averaging kernel for total columns is almost ideal for such profiles and we can use the total columns measured and retrieved at the UNAM campus as a true total column for this purpose.

The IASI-FORLI CO retrieval (Fast Optimal Retrievals on Layers for IASI-CO) shows decreased sensitivity for the lowest layer (George et al., 2009; Turquety et al., 2009; Yurganov et al., 2011) which has to be either compensated for or the IASI-FORLI averaging kernel has to be used to calculate the true column which is expected to be seen by the IASI-FORLI retrieval. The correction of the IASI-columns, shown in the 2^{nd} column of tables 2 and 3, was done with the assumption of a profile with a constant VMR within the mixing layer using the MLH available from each ground-based measurement at UNAM.

$$TC_{IASI}^{corr} = \left(\frac{\int AKT_{IASI} \cdot prf_2(MLH)dz}{\int prf_2(MLH)dz} \right)^{-1} (TC_{IASI} - TC_a) + TC_a \tag{10}$$

The first term in the parenthesis of Eq. 10 is similar as AM in Eq. 6 of the main article, however, for each measurement an actual MLH reconstructed for the time of measurements is used. Its value is $\approx AKT_{IASI}(MLH/2)$ as the CO concentration is assumed to be constant in the mixing layer. Same as for the ground-based CO columns (UNAM), the MLH is averaged over one hour ($\pm 0.5h$) for the dataset without ventilation, Tab. 2, and over half an hour ($\pm$ 15 minutes) for the days including ventilation and weekends, Tab. 3.

Like it is more conventional for the comparison of retrieval with large difference in DOFS, the comparison was done also by degrading the ground-based measurement with the averaging kernel from IASI-FORLI following the equation bellow.

$$TC_{UNAM}^{IASI} = AKT_{IASI}(x_{true}^{UNAM} - x_a) + TC_a \tag{11}$$

The ground based profile retrieval at UNAM has for construction only 1 DOFS (Stremme et al., 2009), but the reconstructed profile x_{true}^{UNAM} is reconstructed from a) the CO column measurement and b) the CO surface concentration of the 5 nearest insitu measurements, so that the profile shape described by the mixing layer height and mixing layer concentration is reliable and equation 11 can be applied. The results are presented in the 3^{rd} column of tables 2 and 3. The 4^{th} column in both tables shows the statistical results of this second comparison.

Table 2. Direct comparison of ground-based UNAM with IASI columns at 10:19 LT, weekdays with low ventilation

Coincidence $\Delta t < 30min$		Total Column (true, UNAM) Equation 10				Total Column (IASI) Equation 11				linear relation $\Delta TC_{IASI} = f(\Delta TC_{UNAM})$			Corr. coef.
										$y = \alpha x$	$y = \beta + \alpha x$		
Radius	days	UNAM	IASI	diff.	err.	UNAM	IASI	diff.	err.	slope	offset	slope	Pearson
km	N	$\times 10^{18} [molec./cm^2]$		%		$\times 10^{18} [molec./cm^2]$		%		-	$[10^{18}/cm^2]$	-	R
8	2	2.22±0.25	2.05±0.43	7	22	1.64±0.09	1.55±0.02	5	5	-	-	-	-
9	5	2.34±0.10	2.63±0.56	-12	24	1.78±0.08	1.93±0.31	-8	18	1.31±1.0	-4.67±1.1	3.11±5.9	0.6
10	6	2.35±0.08	2.86±0.47	-21	20	1.71±0.07	1.93±0.25	-13	15	1.40±0.7	-4.85±0.8	3.28±4.8	0.6
11	8	2.37±0.09	2.64±0.37	-11	15	1.74±0.07	1.85±0.19	-6	11	1.13±0.5	-0.15±0.7	1.18±3.1	0.3
12	10	2.43±0.09	2.57±0.29	-5	12	1.80±0.07	1.84±0.15	-2	9	1.00±0.4	1.29±0.5	0.52±2.2	0.2
13	12	2.41±0.07	2.39±0.27	1	11	1.80±0.06	1.75±0.13	2	8	0.89±0.3	0.60±0.5	0.74±2.0	0.2
14	15	2.45±0.06	2.31±0.23	5	9	1.85±0.05	1.76±0.12	4	6	0.84±0.3	-0.13±0.4	1.00±1.7	0.3
15	17	2.52±0.08	2.33±0.20	7	8	1.91±0.07	1.80±0.11	5	7	0.84±0.3	-0.02±0.3	0.93±1.1	0.4
16	22	2.60±0.08	2.51±0.18	3	7	1.90±0.07	1.83±0.10	3	6	0.86±0.2	1.54±0.3	0.37±0.9	0.2
17	23	2.61±0.07	2.46±0.17	5	7	1.90±0.06	1.81±0.10	4	6	0.84±0.2	1.52±0.3	0.36±0.9	0.2
18	25	2.56±0.08	2.42±0.16	5	7	1.90±0.06	1.80±0.09	4	5	0.85±0.2	1.12±0.3	0.51±0.8	0.2
19	26	2.56±0.07	2.47±0.16	3	6	1.88±0.06	1.82±0.09	3	5	0.87±0.2	1.18±0.3	0.50±0.8	0.2
20	28	2.56±0.07	2.46±0.16	3	6	1.87±0.05	1.80±0.08	3	5	0.88±0.2	0.73±0.3	0.68±0.7	0.3
21	28	2.56±0.07	2.46±0.16	3	6	1.87±0.05	1.80±0.08	3	5	0.88±0.2	0.73±0.3	0.68±0.7	0.3
22	31	2.54±0.07	2.47±0.14	2	6	1.86±0.05	1.82±0.08	2	4	0.90±0.2	0.93±0.2	0.61±0.7	0.3
23	34	2.54±0.06	2.49±0.14	1	5	1.87±0.05	1.83±0.07	2	4	0.92±0.1	0.65±0.2	0.72±0.6	0.3
24	37	2.56±0.06	2.39±0.13	6	5	1.94±0.04	1.84±0.07	5	4	0.84±0.2	0.25±0.2	0.83±0.6	0.4
25	37	2.56±0.06	2.30±0.12	10	5	1.94±0.04	1.80±0.07	7	4	0.79±0.1	0.33±0.2	0.77±0.5	0.4
26	38	2.58±0.06	2.30±0.12	11	5	1.96±0.04	1.80±0.06	8	3	0.77±0.1	0.42±0.2	0.73±0.5	0.4
27	41	2.60±0.06	2.23±0.11	14	4	1.95±0.04	1.75±0.06	10	3	0.69±0.1	0.93±0.2	0.50±0.5	0.3
28	44	2.62±0.06	2.16±0.11	17	4	1.94±0.04	1.69±0.06	12	3	0.63±0.1	1.28±0.2	0.34±0.5	0.2
30	44	2.62±0.06	2.15±0.10	18	4	1.93±0.04	1.68±0.05	13	3	0.63±0.1	1.34±0.2	0.31±0.4	0.2
32	46	2.66±0.06	2.09±0.10	21	4	1.94±0.04	1.64±0.05	15	3	0.57±0.1	1.72±0.2	0.14±0.4	0.1
33	46	2.66±0.06	2.06±0.10	22	4	1.95±0.04	1.63±0.05	16	3	0.55±0.1	1.64±0.2	0.16±0.4	0.1
34	47	2.65±0.06	2.01±0.09	24	4	1.94±0.04	1.60±0.04	17	3	0.52±0.1	1.60±0.1	0.15±0.4	0.1
35	48	2.66±0.06	1.99±0.09	25	3	1.95±0.04	1.60±0.04	17	2	0.51±0.1	1.59±0.1	0.15±0.3	0.1
36	49	2.65±0.06	1.94±0.08	26	3	1.95±0.04	1.58±0.04	19	2	0.49±0.1	1.61±0.1	0.12±0.3	0.1
37	50	2.63±0.06	1.91±0.08	27	3	1.94±0.04	1.56±0.04	19	2	0.48±0.1	1.34±0.1	0.22±0.3	0.2
39	51	2.66±0.07	1.90±0.08	28	3	1.97±0.05	1.56±0.04	20	3	0.44±0.1	1.47±0.1	0.16±0.3	0.1

Table 3. Direct comparison of ground-based UNAM with IASI columns at 10:19 LT, all days and all wind conditions

Coincidence $\Delta t < 15min$		Total Column (true, UNAM) $TC_{apr}^{IASI} + AK_{MLH}^{-1}\Delta TC$				Total Column (IASI) $AK_{IASI}(x_{true} - x_{apr}^{IASI}) + TC_{apr}^{IASI}$				linear relation $\Delta TC_{IASI} = f(\Delta TC_{UNAM})$			Corr. coef.
										$y = \alpha x$	$y = \beta + \alpha x$		
Radius	days	UNAM	IASI	diff.	err.	UNAM	IASI	diff.	err.	slope	offset	slope	Pearson
km	N	$\times 10^{18}[molec./cm^2]$		%		$\times 10^{18}[molec./cm^2]$		%		-	$[10^{18}/cm^2]$	-	R
8	7	2.68±0.17	2.27±0.15	15	8	1.97±0.15	1.79±0.16	9	10	0.79±0.3	1.02±0.3	0.47±0.7	0.5
9	11	2.59±0.11	2.49±0.24	3	10	1.96±0.09	1.93±0.15	1	9	0.93±0.3	2.35±0.4	0.05±1.4	0.0
10	13	2.52±0.10	2.58±0.23	-2	9	1.87±0.09	1.89±0.13	0	8	0.98±0.3	1.86±0.4	0.29±1.2	0.1
11	16	2.49±0.09	2.47±0.19	0	8	1.87±0.07	1.84±0.11	1	7	0.93±0.2	1.50±0.3	0.39±1.0	0.2
12	19	2.50±0.08	2.43±0.16	2	7	1.89±0.06	1.83±0.09	2	5	0.90±0.2	1.49±0.3	0.38±0.8	0.2
13	23	2.47±0.08	2.33±0.15	5	6	1.90±0.06	1.80±0.08	5	5	0.84±0.2	1.05±0.2	0.52±0.7	0.3
14	26	2.49±0.07	2.25±0.14	9	6	1.93±0.06	1.80±0.08	7	5	0.80±0.2	0.92±0.2	0.54±0.7	0.3
15	28	2.52±0.07	2.27±0.13	9	5	1.94±0.06	1.80±0.08	7	5	0.80±0.2	0.64±0.2	0.65±0.6	0.4
16	37	2.54±0.06	2.40±0.12	5	5	1.92±0.05	1.83±0.07	4	4	0.85±0.1	1.18±0.2	0.48±0.5	0.3
17	40	2.54±0.06	2.35±0.11	7	4	1.92±0.05	1.80±0.06	6	4	0.82±0.1	1.08±0.2	0.50±0.5	0.3
18	43	2.52±0.06	2.34±0.10	6	4	1.92±0.05	1.80±0.06	5	3	0.83±0.1	0.92±0.2	0.57±0.5	0.3
19	46	2.52±0.05	2.43±0.11	3	5	1.91±0.04	1.83±0.06	3	3	0.88±0.1	0.74±0.2	0.67±0.5	0.3
20	48	2.52±0.05	2.43±0.11	3	4	1.90±0.04	1.82±0.06	3	3	0.88±0.1	0.51±0.2	0.76±0.5	0.4
21	50	2.52±0.05	2.41±0.11	4	4	1.91±0.04	1.82±0.06	4	3	0.86±0.1	0.54±0.2	0.74±0.5	0.3
22	53	2.51±0.05	2.42±0.10	3	4	1.90±0.04	1.83±0.06	3	3	0.88±0.1	0.68±0.2	0.69±0.5	0.3
23	59	2.52±0.05	2.40±0.10	5	4	1.90±0.04	1.81±0.05	4	3	0.85±0.1	0.60±0.2	0.71±0.4	0.3
24	62	2.54±0.05	2.30±0.10	9	4	1.97±0.03	1.82±0.05	7	3	0.78±0.1	0.36±0.1	0.76±0.4	0.4
25	62	2.54±0.05	2.26±0.09	10	4	1.97±0.03	1.80±0.05	8	3	0.76±0.1	0.42±0.1	0.72±0.4	0.4
26	64	2.58±0.05	2.27±0.09	12	3	1.99±0.04	1.80±0.05	9	2	0.74±0.1	0.69±0.1	0.61±0.3	0.4
27	69	2.60±0.05	2.22±0.08	14	3	1.98±0.03	1.76±0.04	10	2	0.68±0.1	1.09±0.1	0.43±0.3	0.3
28	74	2.61±0.05	2.16±0.08	17	3	1.95±0.03	1.71±0.04	12	2	0.65±0.1	1.15±0.1	0.39±0.3	0.3
30	80	2.61±0.05	2.12±0.08	18	3	1.94±0.03	1.68±0.04	13	2	0.64±0.1	1.24±0.1	0.34±0.3	0.2
32	84	2.63±0.05	2.05±0.06	21	3	1.95±0.03	1.65±0.03	15	2	0.59±0.1	1.44±0.1	0.23±0.2	0.2
33	85	2.63±0.05	2.02±0.06	23	3	1.96±0.03	1.63±0.03	16	2	0.57±0.1	1.37±0.1	0.25±0.2	0.2
34	87	2.63±0.05	1.99±0.06	24	2	1.95±0.03	1.61±0.03	17	2	0.55±0.1	1.29±0.1	0.26±0.2	0.2
35	89	2.61±0.05	1.96±0.06	24	2	1.95±0.03	1.60±0.03	17	2	0.53±0.1	1.20±0.1	0.29±0.2	0.3
36	91	2.60±0.05	1.92±0.05	26	2	1.95±0.03	1.58±0.03	18	2	0.51±0.1	1.18±0.1	0.28±0.2	0.3
37	93	2.60±0.05	1.90±0.05	26	2	1.94±0.03	1.57±0.03	19	2	0.50±0.1	1.10±0.1	0.31±0.2	0.3
39	94	2.61±0.05	1.88±0.05	28	2	1.96±0.03	1.56±0.03	20	2	0.46±0.1	1.20±0.1	0.26±0.2	0.3

2.3 Comparison and validation of column measurements using displaced airmasses

A third strategy is presented to compare the satellite measurements with the ground-based measurement using information about the mean wind transport in the city. The motivation is that there is a large inhomogeneity in the spatial distribution of CO as seen in Figure 6 (main article). Since direct comparison of coincident measurements is performed (Sec. 2.2), a large discrepancy is observed in the results when different coincidence criteria (radii, time) are chosen but the differences found in that exercise are not necessarily a bias. This effort, attempts to demonstrate that a wind propagation operator, here called a footprint, can be constructed from surface wind data in order to know where the CO measured by IASI at 10:19 LT (IASI's overflight time), has been prior to reaching UNAM at noon. At the same time, the growth-rate determined at noon can be back extrapolated to 10:19 LT in order to know which average column is expected over UNAM. These "virtual" column amounts, TC_{IASI}^{UNAM} and $TC_{UNAM}^{10:19}$, respectively, are then compared.

2.4 Footprint for total column measured at UNAM around noon

To reconstruct the average distribution of CO transported to the measurement site we use the same wind data, that are used for classifying days with low ventilation. Implicitly, it is assumed that inside the mixing layer the wind speed and direction do not change systematically with altitude. The wind distribution seeks to describe the average propagation of air masses between IASI's overflight time at 10:19 LT and the interval 11:15-13:15 LT, for which a linear growth of the CO column (local emission flux) was found. The probability distribution of the wind is done in three steps using the 622 days classified as low ventilation days in 2007-2009. i) A probability frequency of wind direction (WD) is calculated using wind sectors of 10°. ii) A frequency distribution for the wind speeds (WS) is calculated for each of the 36 sectors. For each data set, two numbers, describing mean and deviation, are calculated and used for describing the frequency distribution. Because the wind speed cannot be negative, a log-normal distribution is assumed and classified by 4 intervals with the limits [0, $\frac{\mu}{\sigma^{*2}}$], [$\frac{\mu}{\sigma^{*2}}$, $\frac{\mu}{\sigma^{*}}$], [$\frac{\mu}{\sigma^{*}}$, $\sigma^{*}\cdot\mu$], [$\sigma^{*}\cdot\mu$, $\sigma^{*2}\cdot\mu$], with the probabilities of 0.045, 0.136, 0.683 and 0.136, respectively. Both values $\mu = exp(avgr(\ln(WS)))$ and $\sigma^{*} = exp(stdev(\ln(WS)))$ are calculated for each sector corresponding to a wind direction. iii) The normalized frequency distributions are multiplied to obtain a 2 dimensional frequency distribution of direction and velocity $f(WD, WS)$. iv) The distribution is multiplied by the time distance and converted to Cartesian coordinates on a fine grid so that the spatial distribution shown in Figure 3 is achieved (the probability for each segment is divided by the area to obtain a corresponding density). v) The footprint as it is shown in Figure 3 is smoothed before it is linearly interpolated to the grid of the CO column distribution retrieved from IASI and again normalized. This footprint describes a propagation kernel for total columns $G_{t,10:19}^{TC}(lat, lon)$. Its use is analogous with the use of an averaging kernel and can be

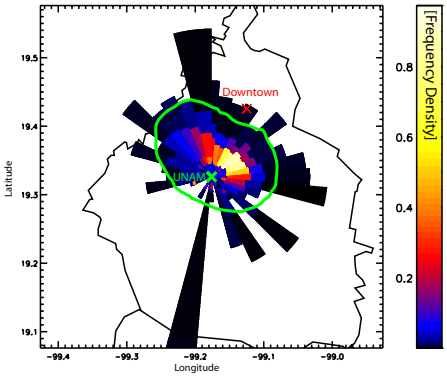

Fig. 3. Probability of wind velocity distribution. The color shows the frequency density of the wind velocities, with its wind-direction given by the angle and its wind-speed given by the radius. The radius is the wind-speed multiplied by the time-distance between UNAM measurements and IASI over-fight time, so that the plot shows where the CO later on measured at UNAM was during IASI overflight. Green area shows where 95% of all the CO measured at 12:00 over UNAM has been at 10:19, when IASI flew over the MCMA.

expressed as follows.

$$TC_{IASI}^{UNAM} = \int G_{t,10:19}^{TC}(lat, lon) \cdot TC_{IASI}(lat, lon) \cdot dlat \cdot dlon \tag{12}$$

TC_{IASI}^{UNAM} is the total column of CO integrated over an area, which is predicted to be measured by the ground-based instrument at UNAM (ideal AKT) around 12:00 LT. This prediction is based on a) the average IASI distribution measured at 10:19 LT, b) the wind data in form of the propagation kernel ($G_{t,10:19}^{TC}$) and c) the false assumption that there has been no emission in this time interval. The propagation kernel $G_{t,10:19}^{TC}$ is shown in Figure 3.

2.5 Virtual coincidence: Total column at IASI overpass

In Section 2.4 the average total column for UNAM at noon based on IASI measurements was predicted with the assumption of no temporal change in the emission. However, as discussed main article Sect.2.2 there is a linear growth of the total column observed of about 0.4×10^{18} molec/cm^2 per hour due to the observed emission flux. Therefore, the column at noon estimated on the basis of IASI might be around 0.7×10^{18} molec/cm^2 higher and might have grown to 3×10^{18} molec/cm^2.

$$TC_{IASI}^{UNAM,12:00} = TC_{IASI}^{UNAM,10:19} + E \cdot (t - 10:19) \tag{13}$$

However, the emission flux is derived from the ground-based measurements and should not be used to manipulate the satellite measurements. Fitting a straight line as it is done in Section 2.2 (main article) with the ground-based total column measurements in the interval 11:15-13:15 LT allows a backwards extrapolation. The total column TC_{IASI}^{UNAM} reconstructed from IASI should not be

Table 4. Intercomparison of mean CO columns at 10:19 LT and around the UNAM campus (2007-2010) estimated from ground (UNAM) and space (IASI) based measurements. The meanvalue is not calculated from the direct averaging of the measurements (see text).

CCA-UNAM (ground based)			IASI-METOP (space-based)		
column	errors[a]		column	errors	
at 10:19	slope	offset	at UNAM	footprint	constraint
$\times 10^{18}$ [molec./cm^2]			$\times 10^{18}$ [molec./cm^2]		
2.38±0.1	± 0.080	± 0.025	2.39±0.1	$< 0.12^b$	$< 0.03^c$

a) errors in the fitted straight line are estimated from the 95% confidence interval

b) from the use of different weighting terms G(t-t'=2h and 3h) as foortprints (see text)

c) from the use of different constraints (clearly underconstrained and clearly overconstrained)

compared with the average column at UNAM at noon, but with $TC^{10:19}_{UNAM}$ which is the total column linearly backwards extrapolated to the time of the IASI overflight time:

$$TC^{10:19}_{UNAM} = TC^{12:15}_{UNAM} - \frac{dTC}{dt} \cdot (t - 10:19) \tag{14}$$

The estimated mean column at 10:19 LT using the columns measurements around noon and its growth rate results in a value $TC^{10:19}_{UNAM}$ of (2.38±0.084)$\times 10^{18}$ molec/cm^2 (error taken from the 95% confidence interval), which is consistent with the IASI measurements TC^{UNAM}_{IASI} of (2.39±0.12)$\times 10^{18}$ molec/cm^2 when the propagation kernel described earlier is used. The results of this comparison is summarized in Tab.4.

References

Rodgers, C. D.: Inverse methods for atmospheric sounding : theory and practice, Ser. Atmos. Oceanic Planet.Phys.,2,World Sci.,Hackensack, N. J., 2000.

Funke, B., López-Puertas, M., García-Comas, M., Stiller, G. P., von Clarmann, T., Höpfner, M., Glatthor, N., Grabowski, U., Kellmann, S., and Linden, A.: Carbon monoxide distributions from the upper troposphere to the mesosphere inferred from 4.7 μm non-local thermal equilibrium emissions measured by MIPAS on Envisat, Atmospheric Chemistry and Physics, 9, 2387–2411, 2009.

Steck, T.: Methods for determining regularization for atmospheric retrieval problems, Appl. Opt., 41, 1788–1797, 2002.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, Journal of Geophysical Research (Atmospheres), 108, 4116, doi:10.1029/2002JD002299, 2003.

Rodgers, C. D.: Characterization and error analysis of profiles retrieved from remote sounding measurements, J. Geophys. Res., 95, 5587–5595, 1990.

von Clarmann, T., Grabowski, U., and Kiefer, M.: On the role of non-random errors in inverse problems in radiative transfer and other applications, J. Quant. Spectrosc. Radiat. Transfer, 71, 39–46, 2001.

Stremme, W., Ortega, I., and Grutter, M.: Using ground-based solar and lunar infrared spectroscopy to study the diurnal trend of carbon monoxide in the Mexico City boundary layer, Atmospheric Chemistry and Physics, 9, 8061–8078, 2009.

George, M., Clerbaux, C., Hurtmans, D., Turquety, S., Coheur, P.-F., Pommier, M., Hadji-Lazaro, J., Edwards, D. P., Worden, H., Luo, M., Rinsland, C., and McMillan, W.: Carbon monoxide distributions from the IASI/METOP mission: evaluation with other space-borne remote sensors, Atmospheric Chemistry and Physics, 9, 8317–8330, 2009.

Yurganov, L. N., Rakitin, V., Dzhola, A., August, T., Fokeeva, E., George, M., Gorchakov, G., Grechko, E., Hannon, S., Karpov, A., Ott, L., Semutnikova, E., Shumsky, R., and Strow, L.: Satellite- and ground-based CO total column observations over 2010 Russian fires: accuracy of top-down estimates based on thermal IR satellite data, Atmospheric Chemistry & Physics, 11, 7925–7942, doi:10.5194/acp-11-7925-2011, 2011.

Turquety, S., Hurtmans, D., Hadji-Lazaro, J., Coheur, P.-F., Clerbaux, C., Josset, D., and Tsamalis, C.: Tracking the emission and transport of pollution from wildfires using the IASI CO retrievals: analysis of the summer 2007 Greek fires, Atmospheric Chemistry and Physics, 9, 4897–4913, doi:10.5194/acp-9-4897-2009, 2009.